

# Opinion: Ethics Plays an Important Role in Artificial Intelligence

By Peter Singer, Project Syndicate, adapted by Newsela staff

05.02.16

<https://newsela.com/articles/ai-ethical/id/16923/>

Last month, AlphaGo, a computer program specially designed to play the game Go, caused shockwaves among aficionados when it defeated Lee Sedol. He is one of the world's top-ranked professional players. He lost the five-game tournament by a score of 4-1.

Why is that news? Nearly 20 years ago, the IBM computer Deep Blue defeated world chess champion Garry Kasparov. We all know computers have improved since then. But Deep Blue won through sheer computing power, using its ability to calculate the outcomes of more moves to a deeper level than even a world champion can. Go is played on a far larger board (19-by-19 squares, compared with 8-by-8 squares for chess). Go has more possible moves than there are atoms in the universe, so raw computing power was unlikely to beat a human with a strong intuitive sense of the best moves.

## Enthusiasm About Artificial Intelligence

Instead, AlphaGo was designed to win by playing a huge number of games against other programs and adopting the strategies that proved successful. You could say that AlphaGo evolved to be the best Go player in the world. It achieved in only two years what natural selection took millions of years to accomplish.

Eric Schmidt, executive chairman of Google's parent company Alphabet, the owner of AlphaGo, is enthusiastic about what artificial intelligence (AI) means for humanity. He said humanity would be the winner, whatever the outcome. Advances in AI will make every human being smarter, more capable and "just better human beings."

Will it? Around the same time as AlphaGo's triumph, Microsoft's "chatbot" – software named Taylor that was designed to respond to messages from people ages 18 to 24 – was having a chastening experience. "Tay," as she called herself, was supposed to be able to learn from the messages she received and gradually improve her ability to conduct engaging conversations. Unfortunately, within 24 hours, people were teaching Tay racist and sexist ideas. When she started saying positive things about Hitler, Microsoft turned her off and deleted her most offensive messages.

## "Smarter Than The Best Human Brains"

I do not know if the people who turned Tay into a racist were themselves racists, or just thought it would be fun to undermine Microsoft's new toy. Either way, AlphaGo's victory and Taylor's defeat occurring at about the same time serves as a warning. It is one thing to unleash AI within a game with specific rules and a clear goal. It is something very different to release AI into the real world, where the unpredictability of the environment may reveal a software error that has disastrous consequences.

Nick Bostrom is director of the Future of Humanity Institute at Oxford University. He argues in his book *Superintelligence* that it will not always be as easy to turn off an intelligent machine as it was to

turn off Tay. He defines superintelligence as an intellect that is “smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills.” Such a system may be able to outsmart our attempts to turn it off.

Some doubt that superintelligence will ever be achieved. Bostrom asked AI experts to indicate dates corresponding to when there is a 1 in 2 chance of machines achieving human-level intelligence and when there is a 9 in 10 chance. The median estimates – or the middle value in a list of numbers – for the 1 in 2 chance were in the 2040-2050 range. For the 9 in 10 chance, it is 2075. Most experts expected that AI would achieve superintelligence within 30 years of achieving human-level intelligence.

## Superintelligence Still A Distant Issue, But Worth Thinking About

We should not take these estimates too seriously. Only 31 percent of the scientists that were surveyed responded. The researchers who did respond work in AI and have an incentive to boost the importance of their field by trumpeting its potential to produce impressive results.

The prospect of AI achieving superintelligence may seem too distant to worry about, especially given more pressing problems. But there is a case to be made for starting to think about how we can design AI to take into account the interests of humans, and indeed of all sentient beings. That would include machines, if they are also conscious beings with interests of their own.

With driverless cars already on California roads, it is not too soon to ask whether we can program a machine to act ethically. As such cars improve, they will save lives, because they will make fewer mistakes than human drivers do. Sometimes, however, they will face a choice between lives. Should they be programmed to swerve to avoid hitting a child running across the road, even if that will put their passengers at risk? What about swerving to avoid a dog? What if the only risk is damage to the car itself, not to the passengers?

## An Ethical Machine

Perhaps there will be lessons to learn as such discussions about driverless cars get started. But driverless cars are not superintelligent beings. Teaching ethics to a machine that is more intelligent than we are, in a wide range of fields, is a far more daunting task.

Bostrom begins Superintelligence with a fable about sparrows who think it would be great to train an owl to help them build their nests and care for their young. So they set out to find an owl egg. One sparrow objects that they should first think about how to tame the owl; but the others are impatient to get the exciting new project underway. They will take on the challenge of training the owl (for example, not to eat sparrows) when they have successfully raised one.

If we want to make an owl that is wise, and not only intelligent, let's not be like those impatient sparrows.

Peter Singer is Professor of Bioethics at Princeton University in New Jersey and Laureate Professor at the University of Melbourne, Australia.

# WEIGHING IN ON THE PROS AND CONS OF ARTIFICIAL INTELLIGENCE

<http://www.buzzle.com/articles/pros-and-cons-of-artificial-intelligence.html>

Artificial intelligence (AI) is the intelligence of machines. It is about designing machines that can think. Researchers also aim at introducing an emotional aspect into them. How will it affect our lives? Read this Buzzle article for an overview of the pros and cons of artificial intelligence.

Artificial intelligence can help lessn the difficulties faced by man, but intelligent machines can never be human.

Intelligence is best defined as the ability of an individual to adapt his/her behavior to new circumstances. Human intelligence is not a single ability but is rather a composition of abilities like learning, reasoning, problem solving, perception, and understanding of language. Since ancient times, people have been thinking of designing machines that will replicate human intelligence. John McCarthy coined the term, 'artificial intelligence' in 1956. He defines artificial intelligence as the science and engineering of making intelligent machines. AI researchers hope to develop intelligent machines, which can perceive, learn and reason like humans. General intelligence is their long-term goal. By general intelligence they mean to incorporate other aspects like social intelligence, judgment, common sense, robotics, and self-awareness into machines. Researchers dream of bringing into machines, factors such as wisdom and the ability to feel, which only humans possess.

## Pros

- ▶ With artificial intelligence, the chances of error are almost nothing and greater precision and accuracy is achieved.
- ▶ Artificial intelligence finds applications in space exploration. Intelligent robots can be used to explore space. They are machines and hence have the ability to endure the hostile environment of the interplanetary space. They can be made to adapt in such a way that planetary atmospheres do not affect their physical state and functioning.
- ▶ Intelligent robots can be programmed to reach the Earth's far-reaches. They can be used to dig for fuels. They can be used for mining purposes. The intelligence of machines can be harnessed for exploring the depths of oceans. These machines can be of use in overcoming the limitations that humans have.

► Intelligent machines can replace human beings in many areas of work. Robots can do certain laborious tasks. Painstaking activities, which have long been carried out by humans can be taken over by the robots. Owing to the intelligence programmed in them, the machines can shoulder greater responsibilities and can be programmed to manage themselves.

► Smartphones are a great example of the application of artificial intelligence. In utilities like predicting what a user is going to type and correcting human errors in spelling, machine intelligence is at work. Applications like Siri that act as personal assistants, GPS and Maps applications that give users the best or the shortest routes to take as well as the traffic and time estimates to reach there, use artificial intelligence. Applications on phones or computers that predict user actions and also make recommendations that suit user choice, are applications of AI. Thus, we see that artificial intelligence has made daily life a lot easier.

► Fraud detection in smart card-based systems is possible with the use of AI. It is also employed by financial institutions and banks to organize and manage records.

► Organizations use avatars that are digital assistants who interact with the users, thus saving the need of human resources.

► Emotions that often intercept rational thinking of a human being are not a hindrance for artificial thinkers. Lacking the emotional side, robots can think logically and take the right decisions. Sentiments are associated with moods that affect human efficiency. This is not the case with machines with artificial intelligence.

► Artificial intelligence can be utilized in carrying out repetitive and time-consuming tasks efficiently.

► Intelligent machines can be employed to do certain dangerous tasks. They can adjust their parameters such as their speed and time, and be made to act quickly, unaffected by factors that affect humans.

► When we play a computer game or operate a computer-controlled bot, we are in fact interacting with artificial intelligence. In a game where the computer plays as our opponent, it is with the help of AI that the machine plans the game moves in response to ours. Thus, gaming is among the most common examples of the advantages of artificial intelligence.

► AI is at work in the medical field too. Algorithms can help the doctors assess patients and their health risks. It can help them know the side effects that various medicines can have. Surgery simulators use machine intelligence in training medical professionals. AI can be used to simulate brain functioning, and thus prove useful in the diagnosis and treatment of neurological problems. As in case of any other field, repetitive or time-consuming tasks can be managed through the application of artificial intelligence.

► Robotic pets can help patients with depression and also keep them active.

- ▶ Robotic radiosurgery helps achieve precision in the radiation given to tumors, thus reducing the damage to surrounding tissues.
- ▶ The greatest advantage of artificial intelligence is that machines do not require sleep or breaks, and are able to function without stopping. They can continuously perform the same task without getting bored or tired. When employed to carry out dangerous tasks, the risk to human health and safety is reduced.

## **Cons**

- ▶ One of the main disadvantages of artificial intelligence is the cost incurred in the maintenance and repair. Programs need to be updated to suit the changing requirements, and machines need to be made smarter. In case of a breakdown, the cost of repair may be very high. Procedures to restore lost code or data may be time-consuming and costly.
- ▶ An important concern regarding the application of artificial intelligence is about ethics and moral values. Is it ethically correct to create replicas of human beings? Do our moral values allow us to recreate intelligence? Intelligence is a gift of nature. It may not be right to install it into a machine to make it work for our benefit.
- ▶ Machines may be able to store enormous amounts of data, but the storage, access, and retrieval is not as effective as in case of the human brain. They may be able to perform repetitive tasks for long, but they do not get better with experience, like humans do. They are not able to act any different from what they are programmed to do. Though this is mostly seen as an advantage, it may work the other way, when a situation demands one to act in way different from the usual. Machines may not be as efficient as humans in altering their responses depending on the changing situations.
- ▶ The idea of machines replacing human beings sounds wonderful. It appears to save us from all the pain. But is it really so exciting? Ideas like working wholeheartedly, with a sense of belonging, and with dedication have no existence in the world of artificial intelligence. Imagine robots working in hospitals. Do you picture them showing the care and concern that humans would? Do you think online assistants (avatars) can give the kind of service that a human being would? Concepts such as care, understanding, and togetherness cannot be understood by machines, which is why, how much ever intelligent they become, they will always lack the human touch.
- ▶ Imagine intelligent machines employed in creative fields. Do you think robots can excel or even compete the human mind in creative thinking or originality? Thinking machines lack a creative mind. Human beings are emotional intellectuals. They think and feel. Their feelings guide their thoughts. This is not the case with machines. The intuitive abilities that humans possess, the way humans can judge based on previous knowledge, the inherent abilities that they have, cannot be replicated by machines. Also, machines lack common sense.

► If robots begin to replace humans in every field, it will eventually lead to unemployment. People will be left with nothing to do. So much empty time may result in its destructive use. Thinking machines will govern all the fields and populate the positions that humans occupy, leaving thousands of people jobless.

► Also, due to the reduced need to use their intelligence, lateral thinking and multitasking abilities of humans may diminish. With so much assistance from machines, if humans do not need to use their thinking abilities, these abilities will gradually decline. With the heavy application of artificial intelligence, humans may become overly dependent on machines, losing their mental capacities.

► If the control of machines goes in the wrong hands, it may cause destruction. Machines won't think before acting. Thus, they may be programmed to do the wrong things, or for mass destruction.

► Apart from all these cons of AI, there is a fear of robots superseding humans. Ideally, human beings should continue to be the masters of machines. However, if things turn the other way round, the world will turn into chaos. Intelligent machines may prove to be smarter than us, they might enslave us and start ruling the world.

It should be understood that artificial intelligence has several pros but it has its disadvantages as well. Its benefits and risks should be carefully weighed before employing it for human convenience. Or, in the greed to play God, man may destroy himself.

Read more at Buzzle: <http://www.buzzle.com/articles/pros-and-cons-of-artificial-intelligence.html>